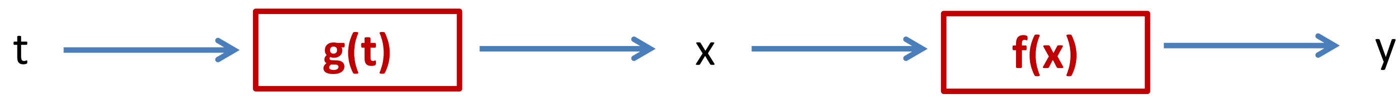
Section 14.5: The Chain Rule (Review)

This is beautiful—but it requires some conceptual thinking. Let's go!

Begin by recalling the 1D chain rule:



Then if we ask "how does the output y change as we change the input t",

Then we are asking for $$\frac{dy}{dt} = \frac{dy}{dx}\frac{dx}{dt}$$

Note the similarity between this and our equation for the Slope of a tangent line

$$f(x) = f(a) + \left.\frac{df}{dx}\right|_a (x-a) \leftarrow \text{Equation for Line tangent to } (\text{a}, f(a))$$

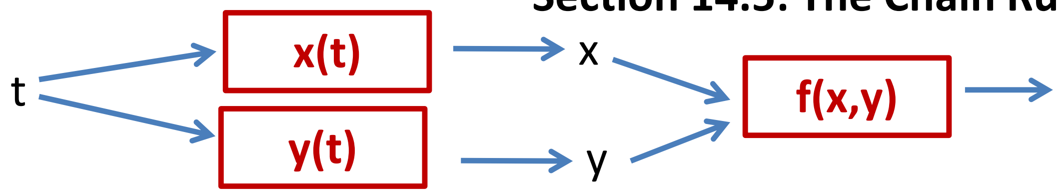
$$f(x) - f(a) = \left.\frac{df}{dx}\right|_a (x-a)$$

Which says that the change in output (f(x)-f(a)) → $\frac{dy}{dt} = \frac{dy}{dx}\frac{dx}{dt}$

equals

the derivative times the change in input (x-a) → $\frac{dx}{dt}$

Section 14.5: The Chain Rule (Review)



$$\frac{dw}{dt} = \frac{\partial w}{\partial x}\frac{dx}{dt} + \frac{\partial w}{\partial y}\frac{dy}{dt}$$

2d: Formula for the tangent plane

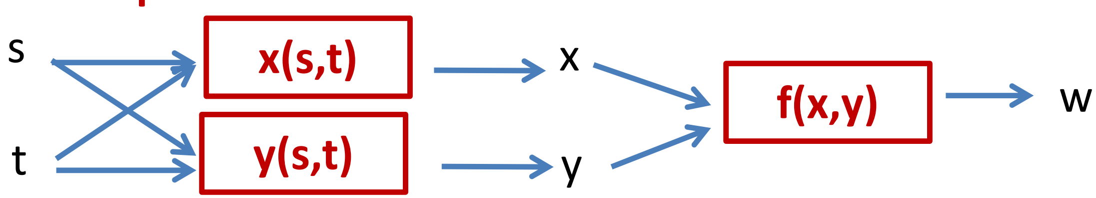
$$f(x,y) = f(a,b) + \left.\frac{\partial f}{\partial x}\right|_{a,b}(x-a) + \left.\frac{\partial f}{\partial y}\right|_{a,b}(y-b) \leftarrow \text{Equation for plane tangent to } (a,b,f(a,b))$$

$$f(x,y) - f(a,b) = \left.\frac{\partial f}{\partial x}\right|_{a,b}(x-a) + \left.\frac{\partial f}{\partial y}\right|_{a,b}(y-b)$$

Which says that the change in output (f(x,y)-f(a,b))
equals
the partial x derivative times the change in x input (x-a)
+
the partial y derivative times the change in y input (y-b)

$$\frac{dw}{dt} = \frac{\partial w}{\partial x}\frac{dx}{dt} + \frac{\partial w}{\partial y}\frac{dy}{dt}$$

A more complicated case



$$\frac{\partial w}{\partial s} = \frac{\partial w}{\partial x}\frac{\partial x}{\partial s} + \frac{\partial w}{\partial y}\frac{\partial y}{\partial s}$$

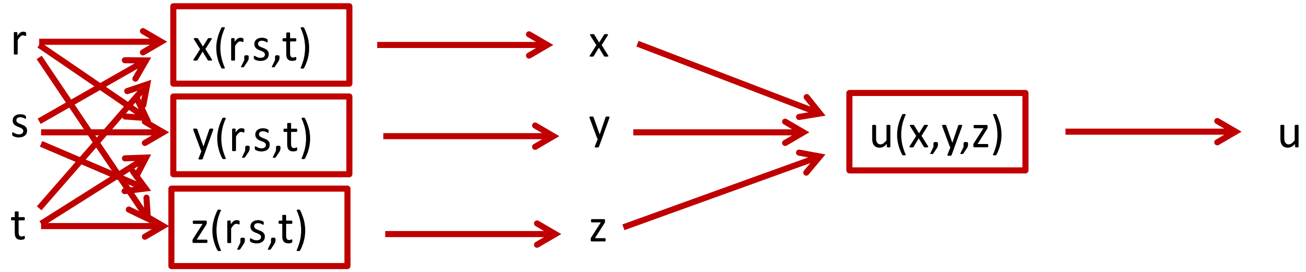
$$\frac{\partial w}{\partial t} = \frac{\partial w}{\partial x}\frac{\partial x}{\partial t} + \frac{\partial w}{\partial y}\frac{\partial y}{\partial t}$$

Section 14.5:

Example: Let $u = x^4y + y^2z^3$ where $x=rse^t$, $y = r s^2 e^{-t}$ and $z = r^2 s \sin t$

Find $\frac{\partial u}{\partial t}$

Answer: Step 1: Make a drawing!!!!!



Step 2: Write the chain of all paths

$$\frac{\partial u}{\partial t} = \frac{\partial u}{\partial x}\frac{\partial x}{\partial t} + \frac{\partial u}{\partial y}\frac{\partial y}{\partial t} + \frac{\partial u}{\partial z}\frac{\partial z}{\partial t}$$

Step 3: Evaluate all the partials

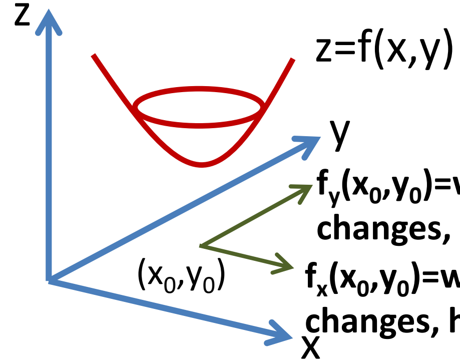
$$= (4x^3y)(rse^t) + (x^4 + 2yz^3)(-rs^2e^{-t}) + (y^2 3z^2)(r^2 s\cos t)$$

Step 4: Substitute to get rid of all the intermediate variables

$$= (4(rse^t)^3(rs^2e^{-t}))(rse^t) + ((rse^t)^4 + 2(rs^2e^{-t})(r^2 s\sin t)^3)(-rs^2e^{-t}) + ((rs^2e^{-t})^2 3(r^2 s\sin t)^2)(r^2 s\cos t)$$

Section 14.6: Directional Derivatives and the Gradient!

We now know what partial derivatives f_x and f_y mean



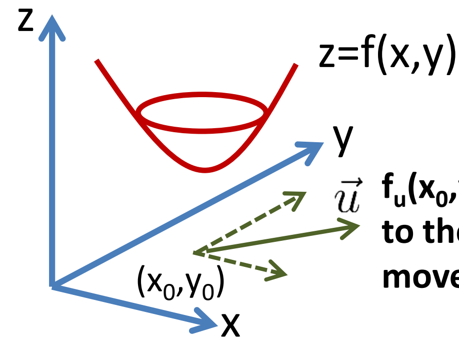
$f_y(x_0,y_0)$=what happens to the output as y changes, holding x fixed

$f_x(x_0,y_0)$=what happens to the output as x changes, holding y fixed

---

But what if you go in a different direction $\vec{u}$? You could ask "how does the output change as you move in the direction u?

This is called the "directional derivative"



$f_u(x_0,y_0)$=what happens to the output as you move in the direction u

Definition:

$$D_{\vec{u}} f(x_0, y_0) \equiv \lim_{h \to 0} \frac{f(x_0+ha, y_0+hb) - f(x_0, y_0)}{h}$$

where $\vec{u}$ is a unit vector with components $\vec{u} = (a, b)$ $\quad |\vec{u}| = 1$

Looks a lot like: $f_x(a,b) \equiv \lim_{h \to 0} \frac{f(a+h,b) - f(a,b)}{h}$ $\quad f_y(a,b) \equiv \lim_{h \to 0} \frac{f(a,b+h) - f(a,b)}{h}$

$\vec{u} = \vec{i} = (a, 0) = (1, 0)$ $\qquad \vec{u} = \vec{j} = (0, b) = (0, 1)$

Section 14.6: Directional Derivatives and the Gradient!

Is there an easy way to calculate:

$$D_{\vec{u}} f(x_0, y_0) \equiv \lim_{h \to 0} \frac{f(x_0+ha, y_0+hb) - f(x_0, y_0)}{h}$$

where u is a unit vector with components $\vec{u} = (a, b) \quad |\vec{u}| = 1$

Yes! And it's so cool!

Claim: $$D_{\vec{u}} f(x_0, y_0) = f_x(x_0, y_0) a + f_y(x_0, y_0) b \quad !!!$$

In other words: to take the directional derivative of f in the direction u, then

$$D_{\vec{u}} f(x_0, y_0) = (f_x, f_y) \cdot \vec{u}$$

Question: why does $\vec{u}$ need to be a unit vector?

Let's use it before proving it: Suppose f(x,y) = $x^3 + 3y^4$. Find $D_u f$ if u = (3/5 , 4/5) at input (2,3)

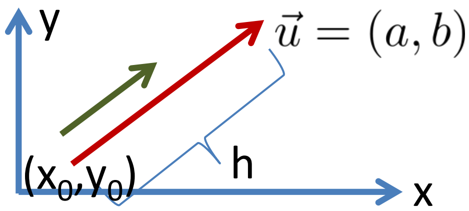
Solution: $$D_u f = (f_x, f_y) \cdot (\tfrac{3}{5}, \tfrac{4}{5}) = (3x^2, 12y^3) \cdot (\tfrac{3}{5}, \tfrac{4}{5}) = 3\tfrac{3}{5}(x^2) + 12\tfrac{4}{5}(y^3)$$

$$D_u f(2,3) = (f_x(2,3), f_y(2,3)) \cdot (\tfrac{3}{5}, \tfrac{4}{5}) = 3\tfrac{3}{5}(2^2) + 12\tfrac{4}{5}(3^3)$$

Let's prove it $D_{\vec{u}} f(x_0, y_0) = (f_x, f_y) \cdot \vec{u}$

Step 1: Let's parameterize by t the green vector leaving (x_0,y_0) and headed in the direction $\vec{u}$

→ $(x(t), y(t)) = (x_0 + ta, y_0 + tb)$



Step 2: We have value for the function f(x,y) at every point on the green vector $f(x(t), y(t)) = f(x_0 + ta, y_0 + tb) \quad ***$

Our strategy: Evaluate the derivative of both sides of * at the input t=0**

Step 3: The expression on the right-hand-side of * is a function of only one variable (namely, t)---since (a,b) and (x_0,y_0) are all fixed, so we can define a new function**

$$g(t) = f(x_0 + ta, y_0 + tb)$$

Step 4: So the derivative of the right-hand-side of g(t) at t= 0 is given by

$$g'(0) = \lim_{t\to 0} \frac{g(t)-g(0)}{t} = \lim_{t\to 0} \frac{f(x_0+ta, y_0+tb) - f(x_0,y_0)}{t} = D_u f(x_0, y_0)$$

The same!

Step 5: On the other hand, we can rewrite the left-hand-side of * as** $g(t) = f(x(t), y(t))$

Step 6: Time for the chain rule: $\frac{dg}{dt} = \frac{\partial f}{\partial x}\frac{dx}{dt} + \frac{\partial f}{\partial y}\frac{dy}{dt} = f_x a + f_y b$

So the derivative of the left-hand-side is $\left.\frac{dg}{dt}\right|_{t=0} = f_x\Big|_{t=0} a + f_y\Big|_{t=0} b = f_x(x_0, y_0)a + f_y(x_0, y_0)b = (f_x, f_y) \cdot \vec{u}$

We now introduce a snappy new notation:

Definition: if f is a function of two variables f(x,y), then we define the gradient $\overrightarrow{\nabla f} \equiv (f_x, f_y)$ $\leftarrow$ **Note: this is a vector**

Example: f(x,y) = sin (xy) Find $\overrightarrow{\nabla f}$

Solution: $\overrightarrow{\nabla f} = (f_x, f_y) = (y\cos(xy), x\cos(xy))$

So now we have a new, nifty notation: $D_u f(x,y) = \overrightarrow{\nabla f} \cdot \vec{u}$

Higher dimensions:

Obviously, we can keep this going to higher dimensions:

if f(x,y,z,w) = $x^3 y^2 + xyzw + w^3$ **find** $\overrightarrow{\nabla f}$ **and find** $D_u f(x,y,z,w) = \overrightarrow{\nabla f} \cdot \vec{u}$ **if u=(1,1,1,2)/√7**

Solution: $\overrightarrow{\nabla f} = (f_x, f_y, f_z, f_w) = (3x^2y^2 + yzw, 2x^3y + xzw, xyw, xyz + 3w^2)$

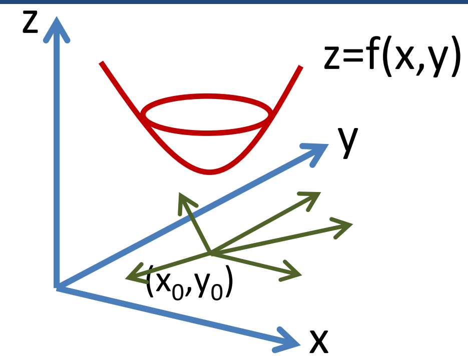
$$D_u f = \overrightarrow{\nabla f} \cdot \vec{u} = [(3x^2y^2 + yzw, 2x^3y + xzw, xyw, xyz + 3w^2)] \cdot \left(\frac{1,1,1,2}{\sqrt{7}}\right)$$

And now comes an amazing moment:

$$D_u f(x,y) = \overrightarrow{\nabla f} \cdot \vec{u}$$

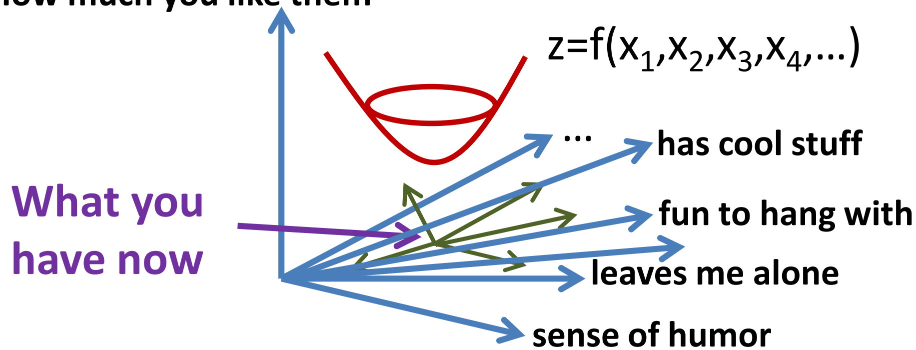
Gives the directional derivative in the direction u

Standing at (x_0,y_0) in which direction should you go to change f the most?



Example: f = how much I like my roommate/brother/sister
= f(sense of humor, leaves me alone, fun to hang with, got cool stuff,….)

If you could pick any unit vector u, what direction would you go maximize how much you like your roommate/brother/sister?



The mathematical version of this question

What direction makes $D_u f(x,y) = \overrightarrow{\nabla f} \cdot \vec{u}$ **the biggest?**

Answer: well,

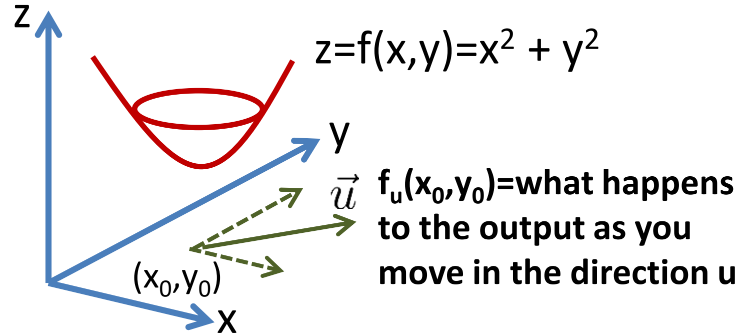
$$D_u f(x,y) = \overrightarrow{\nabla f} \cdot \vec{u} = |\overrightarrow{\nabla f}| \;\; |\vec{u}| \cos\theta$$

Which is biggest when cos θ is biggest which happens when $\overrightarrow{\nabla f}$ and $\vec{u}$

point in the same direction!

Again: $D_u f(x,y) = \overrightarrow{\nabla f} \cdot \vec{u}$ **is biggest when** $\overrightarrow{\nabla f}$ and $\vec{u}$ **point in the same direction!**

Again: standing at (x_0,y_0), you make the biggest change in the output when you move in the direction $\overrightarrow{\nabla f}$ in input space!



Example: suppose $f(x,y) = x^2 + y^2$

Standing at (1,1) the output is f(1,1)=2
What direction should you move to increase f the most?

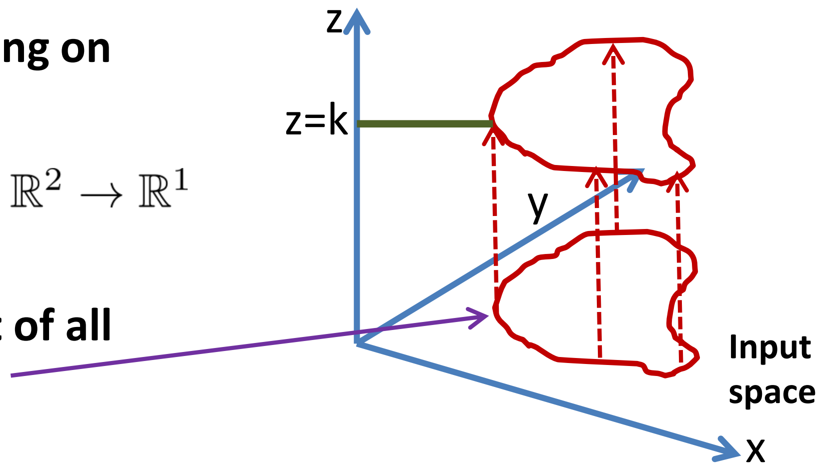
Solution: $\overrightarrow{\nabla f}(x,y) = (f_x, f_y) = (2x, 2y)$ **at input (1,1) this gives the direction (2,2)**

$$D_u f(x,y) = \overrightarrow{\nabla f} \cdot \vec{u}$$

This question helps figure out if you get what is going on

Consider the function f(x,y). This is a mapping from $\mathbb{R}^2 \rightarrow \mathbb{R}^1$

Now, consider the level set k=f(x,y), which is the set of all points in input space that get sent to the output k.

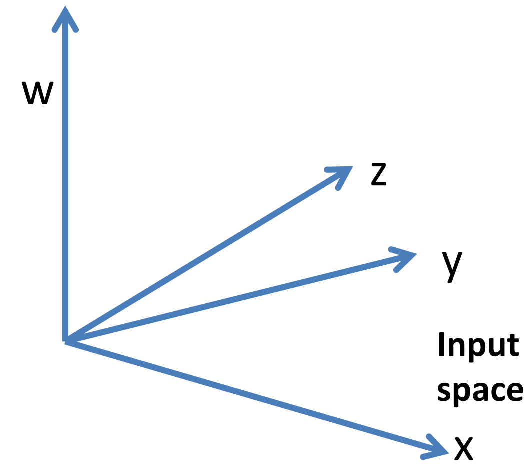


Question: what is the value of the directional derivative $D_u f(x,y)$ as you move along the k level set?

Answer: Zero!!! Because the output does not change as you move on a level set!!!

Let's use these new ideas to think about tangent planes again

$$D_u f(x,y) = \overrightarrow{\nabla f} \cdot \vec{u}$$



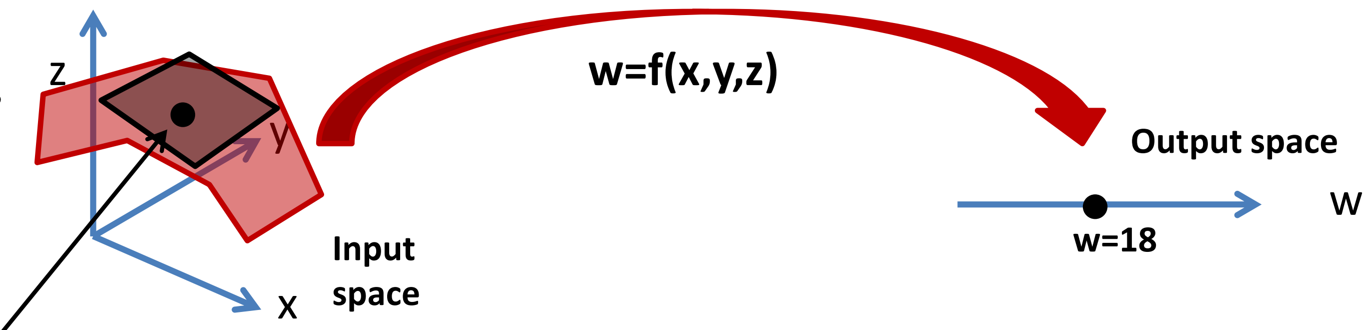
(1) Consider a function of 3 variables: w = f(x,y,z) $\mathbb{R}^3 \rightarrow \mathbb{R}^1$

(2) Requires four dimensions to graph

(3) Let's look at the level set 18 = f(x,y,z) which is the set of all input triples (x,y,z) that get sent to the output w=18

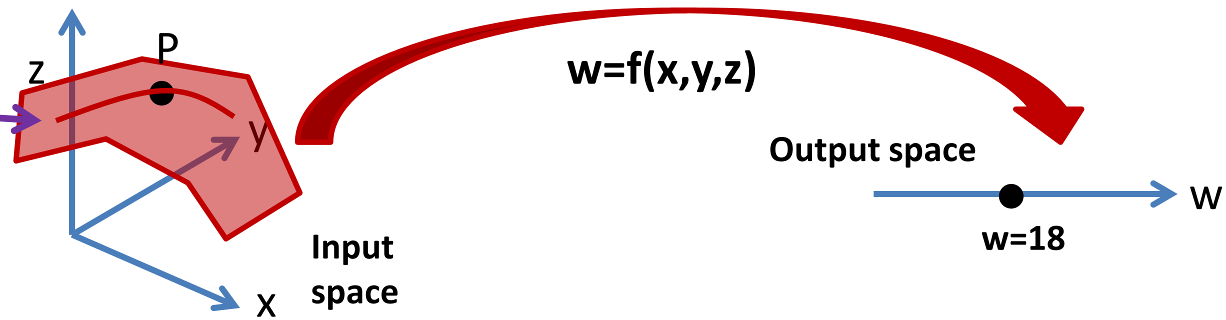
(4) We can draw that surface in input space

Surface of all points in input space sent to 18



(5) Let's find a nice expression for the tangent plane to this level surface at input point $P=(x_0,y_0,z_0)$

Step 1: Imagine a curve (x(t),y(t),z(t)) on level surface: all inputs sent to output 18



So f(x(t),y(t),z(t)) = 18

Step 2: Let's differentiate both sides with respect to t

$$\frac{d}{dt}\left[f(x(t),y(t),z(t))\right] = \frac{d}{dt}[18] \Longrightarrow \frac{\partial f}{\partial x}\frac{dx}{dt} + \frac{\partial f}{\partial y}\frac{dy}{dt} + \frac{\partial f}{\partial z}\frac{dz}{dt} = 0$$

$$\left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z}\right) \cdot \left(\frac{dx}{dt}, \frac{dy}{dt}, \frac{dz}{dt}\right) = 0 \Longrightarrow \overrightarrow{\nabla f} \cdot \vec{r}\,'(t) = 0$$

At t=t_0 $\overrightarrow{\nabla f}(t_0) \cdot \vec{r}\,'(t_0) = 0$

Step 3: $\overrightarrow{\nabla f}(t_0) \cdot \vec{r}\,'(t_0) = 0$ **says that the gradient is perpendicular to every curve in the 18 level set passing through P = (x_0,y_0,z_0)**

Step 4: Ah ha! So $\overrightarrow{\nabla f}(x_0, y_0, z_0)$ **is normal to the tangent plane at the point P=(x_0,y_0,z_0)**

Step 5: Remember that the tangent plane is $a(x - x_0) + b(y - y_0) + c(z - z_0) = 0$

So here is our snappy new formula for a plane tangent to the level surface at the point (x0,y0,z0)

$$\overrightarrow{\nabla f}(x_0, y_0, z_0) \cdot (x - x_0, y - y_0, z - z_0) = 0$$

$$f_x(x_0, y_0, z_0)(x - x_0) + f_y(x_0, y_0, z_0)(y - y_0) + f_z(x_0, y_0, z_0)(z - z_0) = 0$$